Speech Enhancement In Dual Microphones In Mobile Communication Using Adaptive Filters

B.Jessie Jahnavi Rama  
Student of ECE,  
UCEV-JNTUK  
jessiejahnawi@gmail.com

G.Appala Naidu  
Assistant Professor of ECE,  
UCEV-JNTUK  
ganaidu401@gmail.com

Abstract—The speech enhancement is one of the imperative procedures use to enhance the nature of a speech signal i.e. debased by noise. In this paper, a twofold recipient speech change computation for the phones is proposed. The adopted technique exploits the intelligence work calculation and the Kalman channel. This estimation has a clear execution that does not require a forecast of intruding signs estimations. Besides, this count can be used as a piece of little devices with so almost expel between the two recipients Moreover, the usage of such figuring grants reducing diverse commotion sources at various azimuths positions. Finally, the new computation shows its presentations suggesting the perceptual evaluation of talk quality and the time zone waveforms. However the perceptual attributes of the discourse flag relies on the perceptual attributes of human ear. To additionally enhance the execution of discourse upgrade framework, Kalman channel is replaced with wiener channel, which depends on covering attributes of human sound-related framework.

Index Terms—Coherence function, Speech Enhancement, Kalman filter, wiener filter, Peak to signal ratio and Perceptual evaluation of speech quality measurements.

1 INTRODUCTION

SPEECH is one of the main form of communication which is spoken by human. Vocal cord plays one of the major role in supporting the speech communication. The spoken words are created with the help of phonetic combinations. The phonetics consists of limited set of vowels and consonants. Different speech sounds units are found according to the human languages based on the syntax structure and the vocabularies that are used. About 1000 different words vocabularies are available. With the help of the vocal ability speech enables them to sing and even enable to produce different sounds. They are different type of human communications in which speech language and for deaf a gesture form of language is available which is said to be a sign language. Speech is used in different formats, such as in mental processes to enhance and organise the meaning of the interior monologue. Speech also is termed as a basis written language for some cultures.

Speech is examined as far as the Speech creation and speech view of the sounds utilized as a part of vocal dialect. Other research subjects concern speech repetition, the capacity to outline talked words into the vocalizations expected to reproduce them, which assumes a key part in vocabulary development in youngsters and speech blunders. A few scholarly controls concentrate these, including acoustics, brain science, phonetics, subjective science, correspondence studies, and otolaryngology and software engineering. Another territory of research is the means by which the human mind in its diverse ranges, for example, the Broca's region and Wernicke's region underlies speech.

It is disputable how far human discourse is one of a kind; in that creatures likewise speak with vocalizations. While none in the wild have perfectly huge vocabularies,
inquire about upon the nonverbal capacities of dialect prepared primates, for example, Washoe and Kanzi raises the likelihood that they may have these abilities. The birthplaces of discourse are obscure and subject to much level headed discussion and theory.

2LITERATURE REVIEW

Billions of individuals around the globe utilizing cell phones have issues in phone discussion when the discourse is debased by the clamor. Consequently, unique discourse improvement systems were included to lessen the meddling signs in cell phones. Especially, singlemouthpiece commotion diminishment strategies demonstrate its adequacy amid fifty years in many fields as they are anything but difficult to apply [1,2]. Be that as it may, the execution of single channel discourse upgrade calculations is constrained on the non-stationaryclamor case because of the issues identified with the estimation of this commotion utilizing such calculations [2]. Accordingly, discourse quality can be enhanced utilizing multi-channel discourse upgrade calculations [3]. For that, progressively the quantity of mouthpieces is expanded increasingly the nature of discourse will be moved forward. Conversely, an extraordinary number of receivers is hard to actualize in mobile phones and requires more computational many-sided quality. The utilization of double amplifier discourse upgrade calculations can take care of these two issues referred to beforehand. Additionally, these calculations are indicated by its great exhibitions in term of discourse quality and coherence. The writing is advanced by many works which treat a few double channel discourse improvement techniques. Among them, we can refer to the work given by Youssefian et al. [4] which delineated a double channel discourse upgrade calculation utilizing power level distinction for close field. This calculation abuses the distinction of the power motions in the two receivers as a model for clamor decrease. In [5], Youssefian and Lazio exhibited a double amplifier discourse improvement calculation in light of the rationality work. The proposed methodology treats the rationality between the objective and commotion motions as a measure for clamor decrease and can be utilized for barely divided amplifiers. Koldovsky [6] proposed a commotion decrease double receiver in cell phones utilizing a bank of pre-measured target-cancelation channels. This technique depends on an objective cancelation channels misused to assess the commotion, which is then subtracted from the uproarious discourse utilizing Wiener channel or power level contrast calculation. A double receiver discourse improvement strategy in cell phones is proposed in [7]. This technique depends on the between receiver Posteriori SNR Difference (PSNRD) for Speech Presence Probability (SPP) estimation and a MVDR Filter for commotion diminishment. At last, the given work by Prajna et al. [8] exhibited another calculation in light of gravitational hunt calculation (GSA). This approach utilized heuristic calculation for commotion diminishment. Contrasted with this works.

In this paper a dual-channel speech enhancement algorithm dedicated to mobile phone applications using the coherence function and the wiener filter. The proposed method can be applied to all mobile phones as shown in
3 PROPOSED SYSTEM

3.1 Determining Coherence function

Coherence function is presented for the noisy signals. It is one of the parameter which is used to predict the noisy signals. In our proposed work we assume two microphone which are been placed in an environment which is effected with noise. The signals received by the microphones are given as:

\[ L_1(x) = s_1(x) + n_1(x) \]  
\[ L_2(x) = s_2(x) + n_2(x) \]

where

\( L_1(x) \) and \( L_2(x) \) - Noisy signals of the two microphones  
'x' is the simple index,  
\( s_1(x) \) and \( s_2(x) \) - signals obtained at each microphone  
\( n_1(x) \) and \( n_2(x) \) - the noise signal of the two microphones.

The Fourier transform of the two noisy signals can be described as follows:

\[ L_1(y,f) = S_1(y,f) + N_1(y,f) \]  
\[ L_2(y,f) = S_2(y,f) + N_2(y,f) \]

where,

'\( y \)' represents the frame index  
'f' is frequency bin.

The coherence function algorithm is a basic dual microphone speech enhancement method proposed in [9,10]. This technique is based on a correlation between speech signals in the two microphones instead of the noise signals are uncorrelated. The coherence function between the signals \( x_1 \) and \( x_2 \) received by the two microphones is presented as follows:

\[ F_{coh}(y,f) = \frac{P_{x_1x_2}(y,f)}{\sqrt{P_{x_1}(y,f)P_{x_2}(y,f)}} \]  

where,

\( P_{x_1x_2}(y,f) \) - cross power spectral density (CPSD) of the two noisy signals \( x_1 \) and \( x_2 \),  
\( P_{x_1}(y,f) \) and \( P_{x_2}(y,f) \) - power spectral density (PSD) of \( x_1 \) and \( x_2 \).

The goal of the intelligence procedure is to decide whether the speech signal at a particular recurrence container is available or absent though it is relative to the size of the coherence function. The speech signal is administering when the greatness is near one and the noise signal is administering when the extent is near zero. The last theory is genuine when the noise signals at the two receivers are not all that quite connected. Usually, noise and distance are inversely proportional to each other, hence the correlation of the noise signals at the two microphones increases when the distance between them decreases [14].

Thus, the two microphones are too much coherent in our work because they are so closely spaced. The design of the two microphones with target sound sources and the approximate coherence function adopted by this work are detailed in [5]. The two microphones are placed on a duplicate head with a 2-cm distance between them. The clean speech source is placed at 0_ azimuth and the noise source is at b_ azimuth. The distance between the two sources and the microphones are 1.2 m. Based on the previous assumptions, the near coherence function is computed as:

\[ F_{x_1x_2} = \frac{[\cos(\omega t) + j \sin(\omega t)]^{\frac{SNR}{1+SNR}} + [\cos(\omega t \cos \beta) + j \sin(\omega t \cos \beta)]^{\frac{1}{1+SNR}}}{\sqrt{P_{x_1}(y,f)P_{x_2}(y,f)}} \]  

3.2 Speech Enhancement Algorithm

The proposed double channel Speech upgrade algorithm comprises of two stages. The initial step is the double amplifier speech improvement algorithm in view of the coherence function proposed in [5]. The second step is the Kalman channel which displays a decent answer for the direct MMSEtroubles for the stochastic framework by getting an ideal forecast of the perfect discourse [11]. Fig.2. presents the piece graph of the existing speech upgrade calculation.
It demonstrates the time postpone remuneration connected by numerous double microphone speech improvement strategies. In addition, the figure depicts the inexact coherence function characterized by (6).

3.3 wiener filter

Typical deterministic filters are designed for a desired frequency response. However, the design of the Wiener filter takes a different approach. One is assumed to have knowledge of the spectral properties of the original signal and the noise, and one seeks the linear time-invariant filter whose output would come as close to the original signal as possible. Wiener filters are characterized by the following:

3.3.1 Assumption
Signal and (additive) noise are stationary linear stochastic processes with known spectral characteristics (or) known autocorrelation and cross-correlation.

3.3.2 Requirement
The filter must be physically realizable/ causal (this requirement can be dropped, resulting in a non-causal solution) performance criterion: minimum mean square error (MMSE).

3 RESULTS

4.1 Kalman Results
Using the Kalman Filter for the existing algorithm for the good speech quality.

![Fig.2.Existing Algorithm](image1)
![Fig.3.Proposed Algorithm](image2)
![Fig.4.Clean speech signal](image3)
Peak signal to noise ratio and the Perceptual evaluation of speech quality measures for the existing algorithm is

\[
\begin{array}{c|c|c|c}
\text{PSNR} & \text{PESQ for noise signal (dB)} & \text{PESQ for Coherence (dB)} & \text{PESQ for Kalman Filter (dB)} \\
\hline
\text{near signal} & 1.08 & 1.232 & 1.714 \\
\text{coherence} & 27.54 & 27.61 & 28.27 \\
\text{kalman} & 27.2 & 27.4 & 27.6 \\
\end{array}
\]

4.3 Wiener Results
Work extended by using wiener filter for
Obtained better speech signal compared to Kalman filter.

Fig.12. Clean speech signal

Fig.13. Speech signal which is recorded near to microphone

Fig.14. Speech signal which is recorded far from microphone

Fig.15. Coherence enhanced signal

Fig.16. Enhanced signal using wiener filter

4.4 PSNR and PESQ values for Wiener Filter

The below one is output obtained for wiener filter by psnr and pesq values which are run in matlab2014.a.

```
Command Window
PSNR for noise signal dB = 1.080
PSNR for Coherence = 1.232
PSNR for Wiener = 3.044
PSNR of Noise = +27.61 dB
PSNR of Coherence = +27.61 dB
PSNR of Wiener = +30.55 dB
Press 'Enter' button.
Press 'Enter' button.
Press 'Enter' button.
Press 'Enter' button.
Processing ... 100% Done
Successfully Enhanced
>>
```
The comparison between the Kalman and Wiener filter by using the Perceptual Evaluation of Speech Quality and peak-signal to noise ratio values for 30 individual signals are taken restaurant noise at (0db,5db,clear signal).

### TABLE 1

<table>
<thead>
<tr>
<th>Name of the signal</th>
<th>Kalman filter</th>
<th>Wiener filter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>PESQ</td>
</tr>
<tr>
<td>1. sp01</td>
<td>+29.27 dB</td>
<td>1.482</td>
</tr>
<tr>
<td>2. sp01_restaurant_sn0</td>
<td>+28.79 dB</td>
<td>1.118</td>
</tr>
<tr>
<td>3. sp01_restaurant_sn5</td>
<td>+29.03 dB</td>
<td>1871</td>
</tr>
<tr>
<td>4. sp02</td>
<td>+28.04 dB</td>
<td>1.871</td>
</tr>
<tr>
<td>5. sp02_restaurant_sn0</td>
<td>+28.72 dB</td>
<td>1.434</td>
</tr>
<tr>
<td>7. sp03</td>
<td>+27.49 dB</td>
<td>1.420</td>
</tr>
<tr>
<td>8. sp03_restaurant_sn0</td>
<td>+27.91 dB</td>
<td>1.074</td>
</tr>
</tbody>
</table>

### COMPARISON GRAPHS

Comparison graphs between the methods of identification rate comparison of different
processing method of peak signal to noise ratio and mean square error comparison by Spectral subtraction, Normalised least mean squares filter, Recursive least squares, affine projection algorithm, General Kalman Filter for (Speaker Identification) [1] and the speech enhancement using adaptive filters of Kalman and wiener filter by Psnr and Pesq measures.

6 CONCLUSION
The project gives us an improved dual-microphone speech enhancement algorithm for mobile communications. In this system the distance between the two microphones are very much placed closely. So the proposed technique can be used in small mobile phones. First, we depicted the approximate coherence function used in this algorithm. Then, we explained the determined method replacing the Kalman filter with wiener filter. The results obtained with the help of wienerfilter are good compared to the Kalman filter results. Finally, the ability of implementation and intelligibility advantage make this method a good way for future commercial mobile phones.

REFERENCES


